

On the function and the probable origin of Ptolemy's equant

James Evans

61 Boulevard Brune, 75014 Paris, France

(Received 12 September 1983; accepted for publication 18 April 1984)

We examine the roles of the equant point and the eccentric deferent circle in Ptolemy's planetary theory. The necessity of these features of the model is demonstrated empirically by a study of the retrograde arcs of Mars. Finally, these results are used, together with available historical evidence, to suggest the line of investigation that probably led Ptolemy to his discovery of the equant.

I. INTRODUCTION

Most readers are no doubt familiar with the general features of the Greek deferent-and-epicycle theory of the planets. This is a rather remarkable circumstance when one reflects on it: after all, the theory is some two thousand years old and has been obsolete for the last several centuries. But it continues to hold an important place in intellectual history, and in the history of science, as the prototype of physical theory.

The deferent-and-epicycle theory developed over a period of several centuries, roughly centered on the beginning of the Christian era, and was brought into final form by Claudius Ptolemy in the 2nd century A.D. Figure 1 illustrates the model adopted by Ptolemy for Venus, Mars, Jupiter, and Saturn. (Ptolemy's model for Mercury has an extra complication.)

About C as center the deferent circle $AK\Pi$ is drawn. The Earth is at O . Thus the deferent is not centered on the Earth, but is slightly eccentric to it. The line through O and C , when extended, cuts the deferent at the apogee A and the perigee Π of the deferent. The epicycle's center K moves on the deferent, but its motion is not uniform either as seen from the Earth O , or as seen from the center C of the deferent. Rather the motion is uniform as seen from a third center E , the center of uniform motion or *equant point*. That is, an imaginary observer at E would see K travel through equal angles in equal times, while observers at O or C would not. As seen from the Earth, K appears to travel most slowly when it is near the apogee of the deferent, and most rapidly when it is near the perigee. The reference line ET points towards the (effectively infinitely distant) vernal equinoctial point. Thus, according to the definition of the equant, angle $\bar{\lambda}$ increases uniformly with time. The planet P moves on the epicycle at a uniform speed. That is, angle $\bar{\mu}$ also increases uniformly with time.

Popular writers on the history of astronomy have often been unsympathetic toward Ptolemy and his planetary theory. In the melodramatic style that is so frequently cultivated, Ptolemy and his intellectual grandfather, Aristotle, are treated as villains in a story whose heroes are Copernicus, Brahe, and Kepler. The modern chroniclers of science as a combat between adversaries have sought to denigrate Ptolemy by ridiculing his system. Dozens of books could be cited; one example will stand for all. After describing Ptolemy's combination of epicycle and deferent and the three centers—Earth, center of the deferent, and equant point—Rudolf Thiel concedes that the system was successful as a mathematical model. He then begins to register his complaints:

But what happened to the meaning of the whole, to the

Greek cosmology? The planets now traveled in loops, that is to say, around an imaginary point that for unknown reasons itself revolved around the Earth. But even this imaginary point did not revolve precisely around the Earth; the center of its circle was a second imaginary point alongside the Earth. Since this, too, did not yet result in uniform motion, a third imaginary point was needed, from which the motion *appeared* to be uniform!

Such was Ptolemy's ultimate picture; such was the last word in Greek astronomy! Sheer nonsense! What a capitulation for the human mind to accept this sort of mechanics of the universe. What a wretched end for the harmony of the spheres.¹

This is an extreme case of a rather common attitude, which continues to exist despite the remonstrances of historians of science. The prevalence of this view results, at least partly, from the real difficulty of obtaining a sound understanding of the functions served by the various features of Ptolemy's model. In this respect the eccentric deferent and, especially, the equant point are the most troublesome. It is obvious that the epicycle is supposed to account for retrograde motion; the function of the equant is not nearly so transparent. In reality, however, all the features of Ptolemy's model grow in a very natural way out of considerations of the planets' retrogradations.

The present article is divided into three main sections. In Sec. II we give a brief introduction to the phenomena that must be explained by a planetary theory. In Sec. III we demonstrate the empirical necessity of the full Ptolemaic

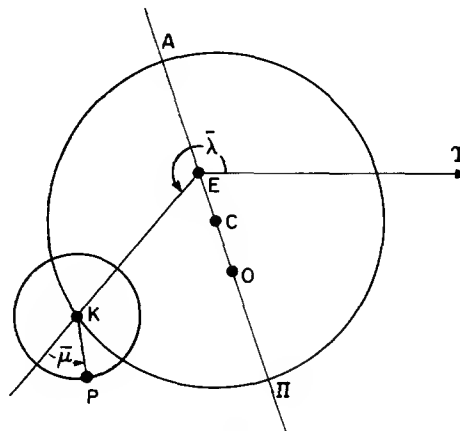


Fig. 1. Ptolemy's theory of longitudes for Venus, Mars, Jupiter, and Saturn.

model. Here, we develop the various features of the model one-by-one and show the role played by each in accounting for certain features of the planets' motions. In particular, we show that the need for the equant is very plainly and strikingly revealed in the pattern of the retrogradations. In Sec. IV we apply the results of this development to the problem of the historical origin of the equant point. Arguing from the available historical evidence, we attempt to reconstruct the line of investigation that led Ptolemy to the discovery of the equant.

II. INTRODUCTION TO THE PHENOMENA

Any departure of a heavenly body from uniform angular speed is called an inequality or an anomaly. This terminology reflects the ordinary manner of calculating a body's position from theory: one proceeds by first calculating the body's mean position—i.e., the position it would occupy in the absence of any irregularity—and then adding to or subtracting from this a small correction or "equation." Ptolemy adopted this plan for the tables in the *Almagest* just as Hipparchos had for his own tables some 250 years before. It still remains the most convenient arrangement.

A. The zodiacal inequality in the case of the sun

Each planet suffers from two quite distinct inequalities, which are intertwined with one another in the sky and which we shall address presently. In the case of the sun, however, there is but a single inequality. A brief examination of the Greek solar theory will serve, therefore, as a good introduction to the theory of the planets, whose motions are more complicated.

The single inequality in the motion of the sun is sometimes called the zodiacal inequality (or anomaly) since it is manifested by the sun's appearing to move more quickly in some parts of the zodiac and more slowly in other parts. For example, the sun's slowest motion occurs at a definite longitude on the ecliptic and at a definite time of year (July, in the present era).

In antiquity, the evidence for a solar inequality rested upon the observed lengths of the seasons: they are not all the same. The seasons are, of course, defined astronomically by means of solstices and equinoxes. In a typical year of the modern era, the equinoxes and solstices fall around these dates: vernal equinox, March 21; summer solstice, June 22; autumnal equinox, Sept. 23; winter solstice, Dec. 22. By simply counting days on a calendar, one finds that the seasons have roughly the following lengths: spring, about 93 days; summer, about 93 days; autumn, about 90 days; winter, about 89 days. Thus it is about seven days longer from the vernal equinox to the autumnal equinox than it is from the autumnal equinox to the following vernal equinox.

The differences in the lengths of the seasons are fairly large and were noticed as early as 430 B.C. by Meton and Euktemon. Reasonably accurate measurements of the lengths of the seasons were not made until somewhat later. By 330 B.C. the lengths, due to Kallippos, were correct to the nearest day. The definitive values for the lengths of the seasons in antiquity were those of Hipparchos, measured around 130 B.C. Hipparchos' values based upon his own observations of equinoxes and solstices, were as follows: spring, $94\frac{1}{2}$ days; summer, $92\frac{1}{2}$ days; autumn, $88\frac{1}{8}$ days; winter, $90\frac{1}{8}$ days. (Note that these are not quite the same as the

lengths of the seasons today. This is not a mistake on the part of Hipparchos: the seasons really have changed in length, although they still add up to the same total of $365\frac{1}{4}$ days. Thus, in antiquity, autumn was the shortest season, but winter is the shortest today.)

Evidently, the sun does not travel at the same angular speed everywhere on its orbital circle. In the modern era it requires 93 days to travel the 90° from the summer solstice to the autumnal equinox, and only 89 days for the 90° from the winter solstice to the vernal equinox. It seems, then, that the sun travels a little faster in January than in July. This apparent variation in speed is called the *solar anomaly* or the *solar inequality*. As mentioned above, it is an example of a zodiacal inequality because the sun's speed at a given moment depends only on its position along the zodiac.

To account for the solar inequality, the Greek astronomers had to give up one or more of their original assumptions about the sun's annual motion. These assumptions were three: (1) the sun's orbit is a circle, (2) centered on the Earth, (3) along which the sun travels at a constant speed. We know today that all three assumptions are false: the orbit is really an ellipse; the Earth is not at its center but at one of the foci; and the sun's speed on the ellipse is not constant. But in the 2nd century B.C. it would have been rash to reject completely a model which evidently required only a minor modification. (The modification will be minor because the lengths of the seasons differ from one another by only a little.) It would have been painful to give up either the circular path or the uniform speed, since this would have destroyed the simplicity of the model and greatly complicated the calculation.

Hipparchos showed that the solar inequality can be accounted for by a much less painful change in the model.² Hipparchos still let the sun move on a circle at uniform speed, but the center of the circle was no longer assumed to coincide with the Earth. Rather, the center of the circle was slightly displaced from the Earth (the center of the world). Hence the sun's orbital circle was said to be *eccentric*.

Figure 2 shows the situation for the present era. That is, we attempt to account for the present lengths of the seasons by means of Hipparchos' model. C is the center of the Sun's eccentric circle, and O is the Earth. The equinoctial and solstitial points still are spaced around the sky at 90° inter-

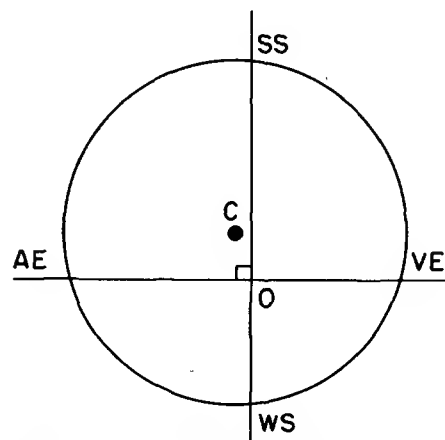


Fig. 2. The solar theory of Hipparchos, applied to save the lengths of the seasons in the twentieth century. VE = vernal equinoctial point; SS = summer solstice; AE = autumnal equinox; WS = winter solstice.

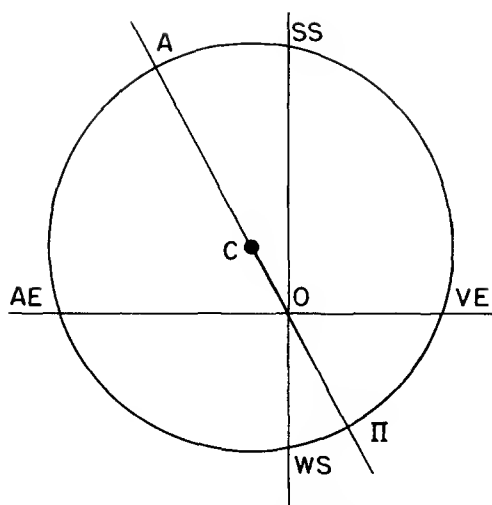


Fig. 3. Hipparchos' solar theory. The line of apsides $A\Pi$ has been added to Fig. 2.

vals, as viewed from the Earth; but these points will no longer divide the circle of the sun's annual motion into equal intervals. Thus we will obtain seasons of unequal lengths. Evidently, the placement of C shown in Fig. 2 produces just the effect desired, for it makes summer the longest season and winter the shortest.

Let us modify Fig. 2 by drawing a line through C and O . This line will cut the circle at two places, as shown in Fig. 3. One of these intersections is the *apogee* of the eccentric circle, marked A . The other is the *perigee* Π . Either A or Π is styled an *apse* or *apsis* ("arch" or "vault") of the orbit. And the line $ACO\Pi$ is called the *line of apsides*. We emphasize that, in Hipparchos' model, the sun travels at a constant speed on its circle. But it will appear to travel more quickly at the perigee and more slowly at the apogee, because of its varying distance from the Earth.

B. The motion of Mars: Two inequalities

In the case of the sun there is a single inequality, the zodiacal inequality, which is manifested by a variation in the sun's angular speed as it moves around the zodiac.

In the case of the planets, in addition to a zodiacal inequality, there is a second inequality which produces the striking phenomenon of retrograde motion. Now, a given planet does not retrogress always at a fixed place on the ecliptic; thus, this dramatic inequality is not an inequality with respect to the ecliptic, but rather is an inequality of a new type. Indeed, a planet's retrogradations are intimately connected, not with its position on the ecliptic, but with its position *relative to the sun*. The superior planets (Mars, Jupiter, and Saturn) always reach the middles of their retrograde arcs when they are in opposition to the sun. The inferior planets (Mercury and Venus) reach the middles of their retrograde arcs when they are in conjunction with the sun. For this reason, the inequality in the motions of the planets that produces retrograde motion is called the *inequality with respect to the sun*.

The zodiacal inequality is sometimes called the *first inequality*. This terminology has its origins in the solar theory: the sun has only one inequality. The additional inequality displayed by all the planets, which causes retrograde motion (i.e., the inequality with respect to the sun), is called, logically enough, the *second inequality*.

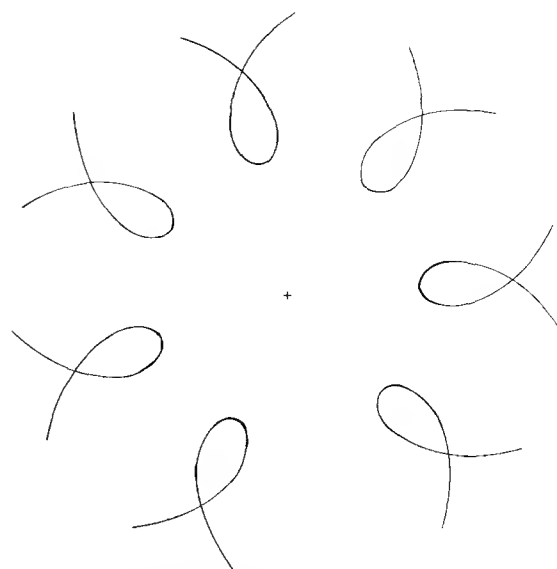


Fig. 4. Retrograde loops of Mars generated by the zero-eccentricity model of Fig. 5.

In the ancient planetary theory the epicycle is the mechanism that produces retrograde motion and is therefore responsible for saving the second inequality. The revolution of the planet on the epicycle produces a series of loops that explain, at least in a general way, the planet's periodic retrogradations. In Fig. 4, typical of illustrations that one often sees in textbooks, all the retrograde loops have the same size and shape, and are uniformly spaced around the ecliptic. However, the model that produces such a pattern is not the complete model of Ptolemy, illustrated in Fig. 1, but rather the simplified model shown in Fig. 5. In this simplified model, the three centers—the equant point, the center of the deferent, and the Earth—all coincide. We shall refer to this as the zero-eccentricity model.

Figure 4 has been produced from the zero-eccentricity model with the use of elements particular to Mars. The radius of the epicycle is 0.66 that of the deferent. The center of the epicycle moves through 114° on the deferent in the time required for the planet to move 100° on the epicycle.

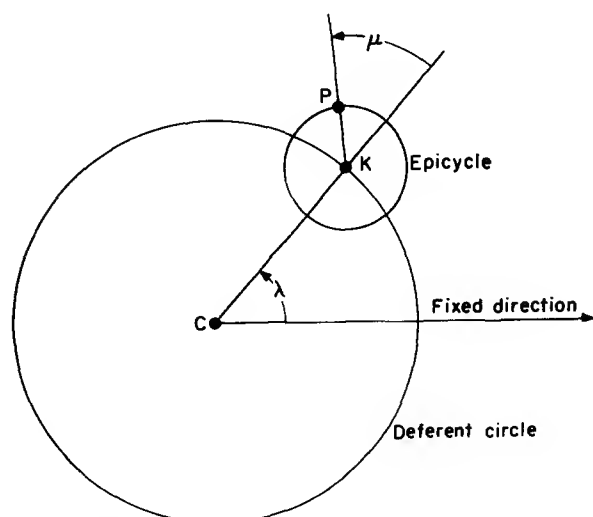


Fig. 5. Zero-eccentricity version of Ptolemy's model.

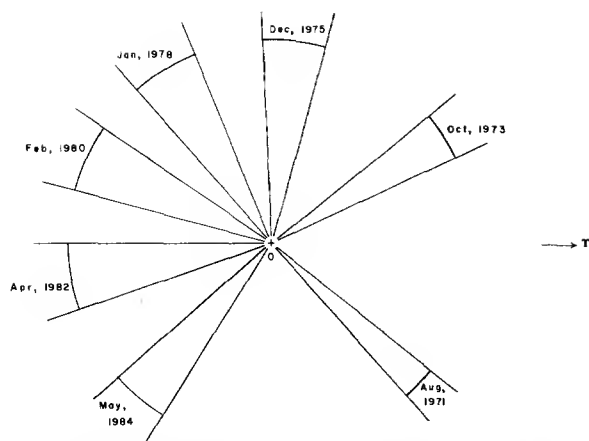


Fig. 6. Retrograde arcs of Mars, 1971–1984. The Earth, from which all observations are made, is marked O. T indicates the direction of the vernal equinox. Longitudes are measured counterclockwise from this direction.

That is, angle $\bar{\lambda}$ increases by 114° while $\bar{\mu}$ increases by 100° . From these relative speeds it follows that, between two successive retrogradations, Mars travels all the way around the ecliptic, plus an arc of about 50° .

The pattern produced by the zero-eccentricity model is simple and elegant. As we have just remarked, all the retrograde loops are the same size and they are separated by uniform intervals of 50° . Thus, if Mars really behaved according to the model of Fig. 5, it would show no zodiacal inequality. This is very far from the true state of affairs.

Fig. 6 shows the actual pattern of the retrogradations of Mars between the years 1971 and 1984. The figure represents the plane of the ecliptic as viewed from its north pole, so longitudes increase in the counterclockwise direction. The Earth is at O. Mars generally travels counterclockwise, but backs up periodically over an arc whose length varies from about 10° (as in 1971) to about 20° (as in 1978). Between one retrogradation and the next, Mars travels all the way around the ecliptic, plus a bit more. The precise distance that the planet travels between one retrogradation and the next is quite variable, with the result that the retrograde arcs are not equally spaced around the zodiac. As may be found by measurement with a protractor on Fig. 6, the centers of the retrograde arcs of 1971 and 1973 are some 75° apart. On the other hand, the centers of the 1978 and 1980 retrograde arcs are only about 34° apart. This is a striking manifestation of Mars' first (or zodiacal) inequality.

How well does the zero-eccentricity model agree with the actual behavior of the planet? Figure 6 illustrates the actual retrograde arcs of Mars for the years 1971 to 1984. If we superimpose Figs. 6 and 4 we obtain Fig. 7. The agreement between theory and observation is strikingly poor. The theoretical plot was constructed by starting Mars in the correct part of the sky for the 1971 retrogradation. The very next retrograde loop, that for 1973, falls well short of the actually observed retrograde arc. Indeed, the positions of nearly all the retrograde loops are quite unsatisfactory. The difficulty is obvious: the uniformly spaced retrograde loops of the model cannot possibly reproduce the unevenly spaced retrograde arcs of the planet itself.

A second difficulty is presented by the uniform width of the theoretical loops. The actually observed retrograde arcs of 1978 and 1980 were about twice as wide as that of

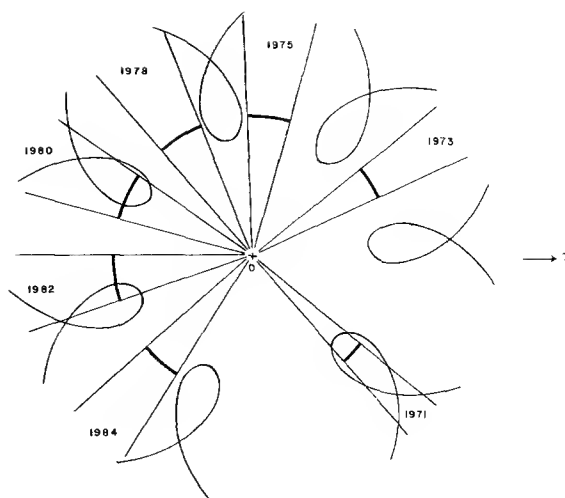


Fig. 7. The retrograde loops predicted by a zero-eccentricity model of Mars superimposed on the planet's actual retrograde arcs. This figure results from the superposition of Fig. 4 and Fig. 6.

1971, and the zero-eccentricity model cannot possibly reproduce this effect. The radius of the epicycle, which is the only adjustable parameter, can be chosen so that the theoretical retrograde loop agrees with, say the *average* size of the observed retrograde arcs. But then in any particular case, the uniform theoretical loop is likely to be either too big or too small. Thus, for example, the 1971 loop is too large and spills over the lines of sight to the planet's actual stations, while the 1978 and 1980 loops are too small.

III. EMPIRICAL NECESSITY OF THE FULL PTOLEMAIC MODEL

We shall now demonstrate how the full Ptolemaic model of Fig. 1 arises in a natural way from the attempt to account for the inequalities in the motions of Mars that are apparent in Fig. 6.

The deficiency of the zero-eccentricity model is that it has no mechanism for producing the inequality with respect to the ecliptic that is so striking in the pattern of Mars' actual retrograde arcs. Mars always retrogrades when it is in opposition to the sun, but the details of the retrogradation depend upon the planet's position on the ecliptic. The retrograde arcs are shortest and farthest apart around longitude 330° (near the position of the 1971 retrogradation); they are longest and closest together around longitude 150° (near the 1978 retrogradation).

A. An intermediate model

Now, as we have seen, the Greek astronomers were forced in the case of the solar theory to adopt a circle eccentric to the Earth, in order to account for the sun's apparent variation in speed along the ecliptic. Perhaps we can save our failing deferent-and-epicycle model for Mars by the same technique. Perhaps the center of Mars' deferent circle is not located at the Earth. In which direction ought we to displace the center of the deferent? If we examine Fig. 6 carefully, we notice a strong line of symmetry, which runs from the 1971 retrograde arc, through the earth O, and on through the gap between the 1978 and 1980 retrograde arcs. The patterns to the left and the right of this line are

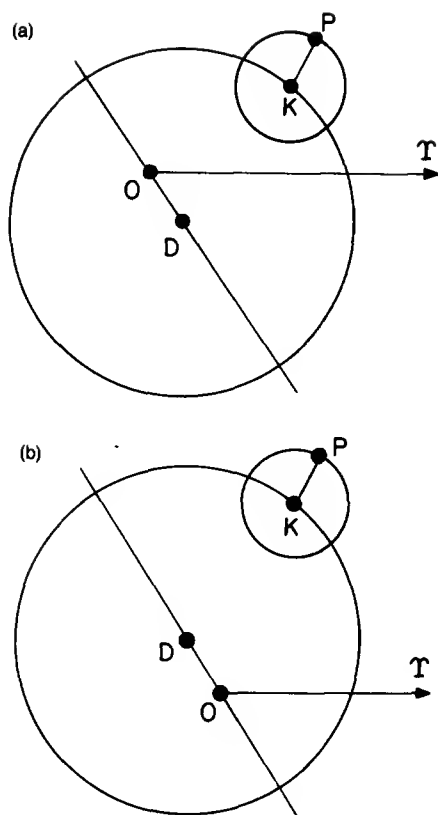


Fig. 8. Two versions of an intermediate model for Mars. There is no separate equant point. But the center D of the deferent may be displaced one way or the other along the line of symmetry in the pattern of the planet's retrogradations. In Fig. 8(a), D has been displaced in the direction of the 1971 retrogradation. In Fig. 8(b), D has been displaced in the direction of the 1978 and 1980 retrogradations.

mirror images of one another. A case could be made, then, for shifting the center of Mars' deferent circle in either of two possible directions; either along the line of symmetry toward the 1971 retrograde arc, or along the line of symmetry toward the 1978 retrograde arc. The result is a model intermediate in complexity between the zero-eccentricity model of Fig. 5 and the final model (Fig. 1) actually adopted by Ptolemy. This intermediate model is illustrated in Fig. 8. The center K of the epicycle travels at a uniform speed around the deferent circle; but the center D of the deferent does not coincide with the Earth O. There are two possible versions of this intermediate model, corresponding to the two possible directions in which D might lie. In Fig. 8(a), the center of the deferent has been located towards the direction of the 1971 retrogradation; in Fig. 8(b), it has been displaced towards the retrogradation of 1978.

What sort of motion results from this intermediate model—in either version? This question is very easy to answer. The retrograde loops will all be of the same size and shape, but the center of the pattern will be at D rather than at O. We may therefore compare the new theory with observation simply by superimposing Fig. 4 on Fig. 6, provided that we shift the center of Fig. 4 with respect to the center of Fig. 6. The center of Fig. 4 is D, the center of Mars' deferent circle; but the center of Fig. 6 is the Earth O, from which we observe the heavens.

Let us first try the version of the intermediate model represented in Fig. 8(a) in which D is located towards the retrograde arc of 1971. The superposition of the theoretical

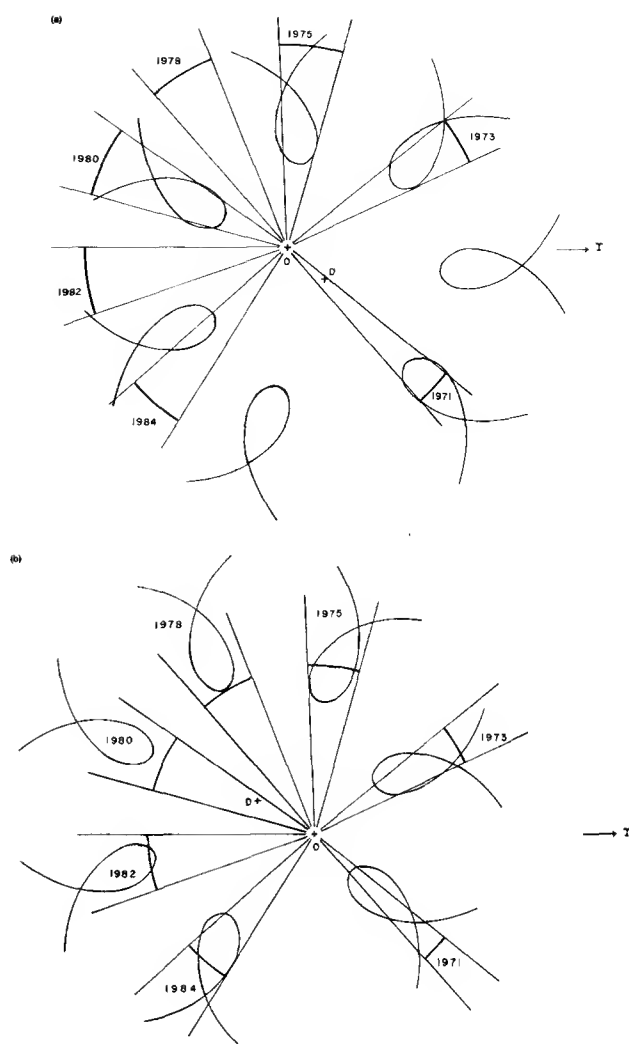


Fig. 9. Comparison of an intermediate model with the real motion of Mars. The Earth is at O, the center of the deferent at D. (a) The model represented in Fig. 8(a). (b) The model represented in Fig. 8(b).

loops with the actually observed retrograde loops results in Fig. 9(a). How well does the new model perform? The *widths* of the arcs are not bad now. Note that both the 1971 and the 1980 loops pretty well fill the space between their lines of sight. Thus, by shifting the center of the deferent towards the 1971 retrogradation, we can make the model produce retrograde loops of about the right width in two opposite parts of the sky. Unfortunately, we have made the *spacing* between the arcs much worse.

Suppose, then, that we examine the motion produced by the model of Fig. 8(b), in which the center of the deferent is located towards the opposite quarter of the sky. The superposition of the theoretical loops and the actual retrograde arcs is given in Fig. 9(b). All of the theoretical loops now fall on top of their corresponding retrograde arcs. Evidently, this version of the intermediate model produces just the kind of variable *spacing* that we need. But what of the *widths*? By shifting the center of the deferent towards the 1980 retrograde loop, we have actually made the widths worse. The 1971 loop is now much too wide, while the 1978 and 1980 loops are much too narrow.

It seems, then, that a simple shift of the deferent's center cannot save our model. In order to make the model reproduce the observed spacing of Mars' retrogradations, we

must locate the center of the deferent in one direction. But in order to produce retrograde arcs of the right widths, we must locate the center of the deferent in the opposite direction. There is no way to produce the correct spacings and the correct widths simultaneously by a simple shift of the deferent's center. This fact can be made even more apparent by the following simple argument. Imagine a row of trees. If we back away from the trees they will appear to become smaller and closer together. But, as Fig. 6 shows, around the direction of the 1971 retrogradation, we need the retrograde loops of Mars to appear small and far apart. There is no way to produce this appearance by simply shifting the position of the Earth with respect to the center of the uniform loops of the zero-eccentricity theory. Some wholly new theoretical device is called for.

B. Ptolemy's final model

Figure 1 illustrates the planetary theory finally adopted by Ptolemy for Venus, Mars, Jupiter, and Saturn, which was described at the beginning of this article.

Before we examine the motion predicted by the new model, let us take a moment to see how it is suggested in the most direct way by the phenomena we are trying to save. We return to Fig. 9(b), which pictures the motion produced by the second version of our intermediate model. In this intermediate model [Fig. 8(b)] the center of the deferent and the center of uniform motion coincide at D, which is slightly eccentric to the Earth O. That is, the center K of the epicycle moves at a uniform speed around the deferent, and the line DK turns at a constant rate. In Fig. 9(b) we have adjusted the position of D (center of the loop pattern) to obtain the best possible agreement with the actually observed positions of the retrograde arcs. As far as the *positions* of the loops go, the intermediate model is excellent: every one of the theoretical loops falls on top of its corresponding observed retrograde arc. The theoretical loops are, of course, equally spaced as seen from D. Since the actual retrograde arcs coincide with them, it follows that D really is the center of the actual planet's uniform motion. That is, *point D is acting as an equant point*. We have no choice in the matter: the planet insists on it.

Of course, the intermediate model of Fig. 8(b) fails rather miserably to produce correct widths for the retrograde arcs. The theoretical loop for 1971 is much too large, while the loops for 1978 and 1980 are much too small. The only apparent solution is precisely that adopted by Ptolemy, i.e., to separate the center of the deferent from the center of uniform motion. In Fig. 8(b), point D represents both the equant and the center of the deferent. We must leave the equant at D, in order to preserve the correct spacing between the retrogradations. But if we move the deferent's center closer to O, this will cause the retrograde loops of 1978 and 1980 to draw nearer the Earth and, therefore, to look larger. The 1971 loop will recede from the Earth and, therefore, look smaller. In this way we will have a fair chance of producing retrograde arcs of correct and variable width, while preserving the correct spacing we have already achieved. The final model, which is produced by the separation of point D into points E and C, is precisely that adopted by Ptolemy and illustrated by Fig. 1.

Of course, it is not obvious how well the new model actually will work. Once the center of uniform motion is separated from the center of the deferent circle, the retrograde loops will cease to be of uniform size and shape. The only

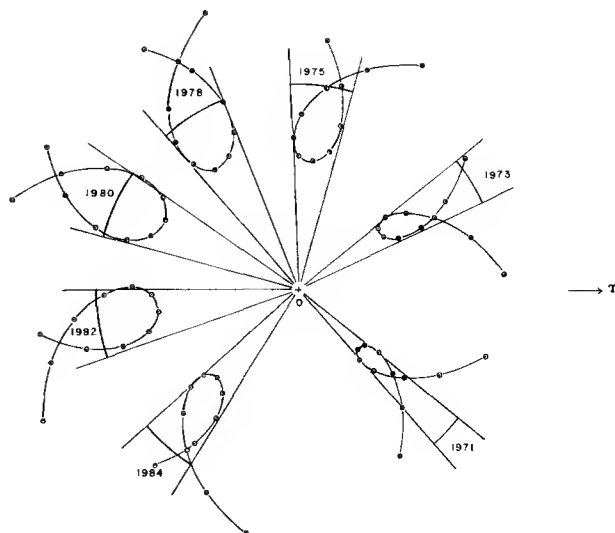


Fig. 10. Comparison of the Ptolemaic theory for Mars with the actual retrograde arcs of the planet between the years 1971 and 1984. The numerical parameters of the model are not Ptolemy's but have been determined anew by applying Ptolemaic procedures to modern data.

recourse is actually to calculate the motion predicted by the model of Fig. 1, using numerical values for the parameters that are appropriate to Mars. (These parameters are the radius of the epicycle, the magnitude and direction of the eccentricity, the rates of motion of the planet around the epicycle and of the epicycle around the deferent, and the positions of K and P at some one particular instant.) If the retrograde loops predicted by the model in Fig. 1 are plotted and superimposed upon Mars' actual retrograde arcs, the result is Fig. 10.

The positions and widths of the retrograde loops in Fig. 10 are clearly excellent. What is perhaps less clear is that the model also saves *the times* at which the retrogradations occur. As remarked above, the superior planets reach the centers of their retrograde arcs very nearly when they are in opposition to the sun. Now, the position of the sun on the zodiac is the measure of time. Thus, if a retrograde loop takes its correct position on the zodiac, the sun, in opposition, also takes its correct position; i.e., the retrogradation occurs at the correct time of year. The final Ptolemaic model can only be judged a stunning success.

In conclusion, it looks as if Ptolemy's model is as simple as the planets themselves will allow. An epicycle is required to account for the second inequality. A simple off-center deferent is not adequate to explain all the manifestations of the first inequality, but the combination of an eccentric deferent with a separate center of uniform motion does very well. The two inequalities are present in the motions of the planets and any planetary theory with any pretensions of validity must include mechanisms for accounting for both.

C. Note on Jupiter and Saturn

The three superior planets (Mars, Jupiter, and Saturn) all behave in generally the same fashion. However, certain aspects of the retrogradations of Jupiter and Saturn set them apart from Mars.

Most conspicuously, the length of the retrograde arc scarcely varies in the cases of Jupiter and Saturn. We have seen that Mars' longest retrograde arc (20°) exceeds its

shortest (10°) by a factor of two. In the case of Jupiter, the lengths of the retrograde arcs at apogee and at perigee differ by only a minute or two of arc—although the average length of the arc is about 10° . Saturn shows only a little more variation. Saturn's longest retrograde arc exceeds its shortest by only a third of a degree, the average retrograde arc being about 7° . The chief manifestation of the zodiacal inequality in the cases of Jupiter and Saturn is therefore the uneven spacing of the retrograde arcs around the zodiac. (The variation in spacing is quite respectable, being fully half that in the case of Mars.³)

In spite of this difference, any attempt to apply the intermediate model to Jupiter or Saturn leads to exactly the same result as it did in the case of Mars. Consider Jupiter, for example. The zero-eccentricity model accounts for the equal-sized retrogradations, but cannot account for the uneven spacing. Version (b) of the intermediate model must therefore be adopted to save the spacing. But, in the process, the retrograde arcs at apogee will be made too small, while those at perigee will be made too big. Again, it will be necessary to separate the center of the deferent from the center of uniform motion, in exactly the same way as before: the center of deferent must lie between the Earth and the center of uniform motion.

IV. ON THE HISTORICAL DEVELOPMENT OF GREEK PLANETARY THEORY

Of the history of planetary theory before Ptolemy, we know very little. The difficulty, of course, is the inadequacy of the sources. Mathematical astronomy was in antiquity, as it remains today, the province of a handful of specialists. After the planetary theory was put into definitive form by Ptolemy, the works of his predecessors lost their scientific interest and ceased, therefore, to be read and copied. Most of what we know of the history of planetary theory before Ptolemy depends upon a few remarks made by Ptolemy himself in the *Almagest*.⁴

In the following pages we shall suggest the probable historical development of the planetary theory down to Ptolemy's discovery of the equant point. We shall argue that this development followed more or less the process that we have just outlined. By this we do not mean that the Greeks drew figures of the retrograde loops like ours. Rather, we assert that it was the desire to save the lengths of the retrograde arcs and the intervals between them, particularly in the case of the superior planets, that led Ptolemy to his final model. Our figures have provided a method of rendering visible all at one time the results of hundreds of calculations. If we have used modern observations and elements in drawing the figures, it was in order to avoid the difficult and largely irrelevant question of the validity and precision of the ancient observations. To forestall a possible objection, it suffices to note that, at least in the case of Mars, the disagreement between the real retrogradations and the zero-eccentricity model, or between the real retrogradations and either version of the intermediate model, is vastly larger than even the most conservative estimate of the observational error of the Greek astronomers—a fact that is evident in the figures themselves. That the development of the planetary theory followed the course outlined above is therefore perfectly possible. It remains to demonstrate, by an examination of the historical evidence, that this was the probable course of development.

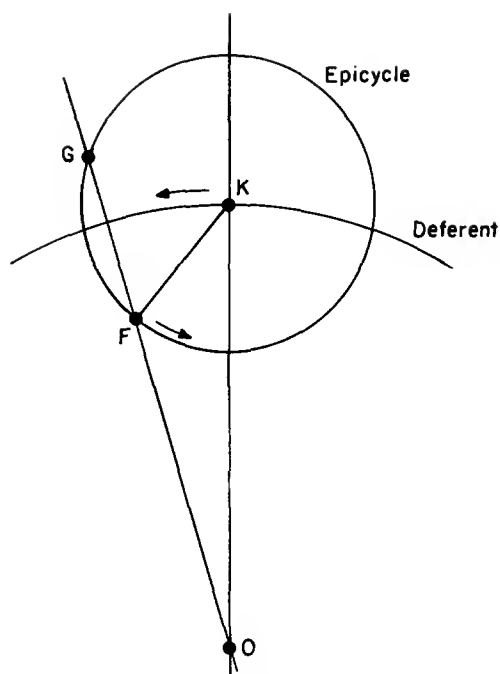


Fig. 11. Illustrating Apollonius' theorem.

A. The zero-eccentricity deferent-and-epicycle model

It appears that the deferent-and-epicycle theory originated in the 3rd century B.C. with the geometrical researches of Apollonios of Perge. In *Almagest* XII, 1, Ptolemy says that Apollonios and the other mathematicians of his school proved a theorem concerning the location of a planet's stationary points. Refer to Fig. 11. Here O represents both the Earth and the center of the planet's deferent circle, while K is the center of the epicycle. Let f_λ denote the angular speed with which K moves about O. (That is, f_λ is the angular speed at which the line OK turns.) Let f_μ denote the angular speed of the planet on the epicycle, referred to the moving line OK. (If the planet is at F, say, then f_μ is the rate at which angle FKO changes.)

Apollonios proved that the planet, at F, will appear stationary, as seen from O, whenever

$$\frac{1}{2}(FG/OF) = f_\lambda/f_\mu.$$

This theorem permits one to calculate the length of the retrograde arc if the two angular speeds and the radius of the epicycle are known. The angular speeds are determined by the planet's tropical and synodic periods. The radius of the epicycle could then be fixed by requiring the model to produce retrograde arcs of some desired length.

Apollonios' theorem, it should be noted, applies directly only to a zero-eccentricity model. The first practitioners of the deferent-and-epicycle model evidently concerned themselves solely with the inequality with respect to the sun, and took no account of the zodiacal inequality.⁵

However, it is very likely that these astronomers were aware of the zodiacal anomaly. As we have seen, in the case of Mars, this inequality is very striking. There is a difference of 10° between the lengths of the shortest and the longest retrograde arcs. The distance between the centers of successive retrograde arcs varies from a minimum of 34° to a maximum of 75° —an enormous variation. Even the crudest attempts at quantitative measurements of the planet's

positions at its stations would have disclosed this inequality. Moreover, Ptolemy tells us, in *Almagest* IX, 2, that the stations were one of the two classes of planetary phenomena of greatest interest to his predecessors, the other being heliacal risings and settings. It seems, therefore, that the school of Apollonios deliberately neglected the zodiacal inequality and attempted only to provide a model that was qualitative and explanatory in nature. The idea that one could demand a quantitative and predictive model must have dawned very slowly.

B. Demise of the zero-eccentricity model and the rise of the intermediate model

By the 2nd century B.C., the Greek astronomers began to grapple with the planets' first inequality. Our chief source of information on this development is Ptolemy's brief summary (in *Almagest* IX, 2) of Hipparchos' work on the planetary theory. As Ptolemy remarks, Hipparchos made notable contributions to the theories of the sun and the moon. However, according to Ptolemy, he made no attempt to give a theory of the planets, but only arranged the observations in a more useful way and showed the appearances to be inconsistent with the hypotheses of the mathematicians. Hipparchos noted that, owing to the two separate inequalities, the retrogradations of each planet are not uniform; but the mathematicians gave their geometrical demonstrations as if there were a single inequality and as if all the retrograde arcs were of the same length. From this remark by Ptolemy, it appears that one part of Hipparchos' contribution was a demonstration that the zero-eccentricity model of his predecessors was inconsistent with the real motions of the planets. The most obvious difficulty was, of course, the uneven spacing and unequal widths of the retrograde arcs, as illustrated in Fig. 6. Again, Hipparchos' immediate predecessors could hardly have been unaware that the planets' retrogradations are unevenly spaced around the zodiac, for this is reflected quite clearly in the unequal times that separate successive retrogradations. And, at least in the case of Mars, the unequal lengths of the retrograde arcs must also have been known. Hipparchos' insistence that a planetary theory ought to account for these inequalities was therefore even more significant for the future of astronomy than was his simple notice that theory had so far failed to make such an accounting.

Ptolemy also says that Hipparchos showed that the appearances had not been saved either by eccentric circles, or by circles concentric with the Earth but bearing epicycles, or even by eccentrics and epicycles together. A model with eccentrics and epicycles together would be something like that illustrated in Fig. 8. Thus it appears that, at least by Hipparchos' time, the intermediate model had also been investigated.

It is even possible to say which version of the intermediate model was preferred. Pliny, in Book II of his *Natural History* (1st century A.D.), has a little to say about the planets. Pliny was not a trained astronomer, and much of his discussion is both confused and confusing. However, his writing predates Ptolemy by two generations and he had access to pre-Ptolemaic works that are now lost. One of Pliny's remarks bears in an important way on the state of planetary theory before Ptolemy. Pliny⁶ remarks that the planets have their circles drawn about different centers. The apogees of the superior planets' deferents, according to Pliny, are as follows: Saturn, in Scorpio; Jupiter, in Virgo;

Mars, in Leo. These are consistent with version (b) of the intermediate model. That is, the apogees were placed correctly to account for the *spacing* of the retrogradations about the zodiac. No account, therefore, was taken of the *lengths* of the retrograde arcs.

Version (b) was indeed, the more reasonable choice. Except in the case of Mars, the variation in the lengths of the retrograde arcs is not very dramatic. For most of the planets, this variation could therefore be ignored, while the *uneven spacing* of the arcs could not. Besides, version (b) of the intermediate model had a close parallel in the solar theory. The center of the planet's epicycle (or the epicycle itself) is analogous to the sun. The correct apparent motion of the epicycle around the deferent was obtained by placing the apogee of the deferent in the quarter of the sky where this motion was apparently slowest.⁷

Hipparchos' suspicion that even eccentric circles and epicycles together were inadequate to explain the appearances very likely was based on a realization that the intermediate model, which could produce the correct spacing, failed to produce retrograde arcs of the correct widths. However, as Hipparchos had no theory to offer in its place, this intermediate model continued in use down to Ptolemy's time.

C. Discovery of the equant

As far as anyone knows, the equant originated with Ptolemy. With this discovery, he brought the theory of longitudes into a highly satisfactory form, which it was long to retain. Unfortunately, it is difficult to be sure of the manner in which he arrived at his discovery. Ptolemy's style in the *Almagest* is the style of most scientific writing. Mathematical propositions are made and proved; observations are reported and used to derive certain parameters; these parameters are then used as the foundation upon which other parts of the theory are based. The style is lean, elegant, and efficient. It also discloses very little of the original process of discovery.

Ptolemy introduces the equant in *Almagest* IX, 5, in the course of a general overview of the theory of longitudes. He says simply that the epicycle's center is carried on a circle (i.e., the deferent) whose center is distinct from the center of uniform motion; moreover, in the case of all the planets but Mercury, the center of the deferent is midway between the center of uniform motion and the Earth. For this assertion, Ptolemy offers no justification. This rather abrupt statement of hypothesis is uncharacteristic of Ptolemy, who usually takes pains to explain the reasons that lie behind his choice of a model.

In *Almagest* IX, 2, Ptolemy had apologized for the fact that he might seem to presuppose things without immediate foundation in the phenomena. He justified himself by saying that things supposed without proof cannot be without some logic if they are found to be consistent with the appearances, even though the way of arriving at them might be hard to explain. It is quite possible that this is a veiled reference to his manner of introducing the equant, which follows shortly afterward.

Thus Ptolemy nowhere says explicitly how he arrived at the idea of the equant, which was perhaps his most important personal contribution to the traditional planetary theory. Consequently, this puzzle has long attracted the attention of astronomers and historians of science.⁸

Now, in his treatment of Venus, Ptolemy gives a method

of determining the position of the equant point by direct calculation (*Almagest* X, 3). Thus, in the case of Venus, Ptolemy *seems* to prove from observational data that, indeed, the center of uniform motion is distinct from the center of the deferent and that the latter is twice as far from the Earth as is the former. Some have believed to see here the original discovery of the equant point.⁹ While this hypothesis cannot be disproved, it is highly improbable, for the following reasons. (1) Before beginning his calculation of the position of Venus' equant, Ptolemy must have already realized that the equant point was distinct from the center of the deferent, or else there could have been no motive for undertaking the calculation. (2) Ptolemy finds, purportedly by direct calculation, that the center of the deferent lies exactly halfway between the Earth and the equant point. But, in fact, Ptolemy's procedure, if applied to accurate data, would not lead quite exactly to this result. Ptolemy's calculated result, therefore, probably did not represent an original discovery, but was rather a confirmation of a hypothesis he had held already. (3) The dates of the latest observations used for each planet tend to suggest, although quite inconclusively, that Ptolemy completed his work on the superior planets before the inferior planets.¹⁰ We must therefore search elsewhere for the origin of the equant point.

What was the course of investigation that led Ptolemy to the equant? It seems most likely that he was experimenting with the two versions of the intermediate model, which he found irreconcilable. In *Almagest* X, 6, where Ptolemy takes up the derivation of the parameters for the superior planets, we find some evidence that this was the case. Ptolemy asserts, without proof, that for these planets, as for Venus, the center of the deferent lies exactly halfway between the Earth and the equant point. And then he says something, by way of justification, that is extremely interesting. Ptolemy says that *the eccentricity calculated from the zodiacal anomaly is about twice the eccentricity calculated from the lengths of the retrograde arcs at greatest and least distances*.

Ptolemy gives no details of his method, but it appears that, for one or more of the superior planets, he calculated the value of the eccentricity required to save the zodiacal anomaly as manifested in the motion of the epicycle's center. That is, he found the distance OD in Fig. 8(b) that would give the correct *spacing* of the retrograde arcs, as in Fig. 9(b). This was, of course, a part of the procedure he had inherited from his immediate predecessors, who subscribed to version (b) of the intermediate model. Ptolemy also calculated the distance OD in Fig. 8(a) that would give the correct *lengths* of the retrograde arcs at two opposite parts of the sky, as in Fig. 9(a). This involved methods based upon Apollonios' theorem. The two calculations of the eccentricities required to save the phenomena, therefore, were both based upon traditional procedures. Ptolemy found that the two results *were not the same*. As shown clearly by a comparison of Figs. 9(a) and 9(b), the eccentricity (distance OD) required to save the spacing [Fig. 9(b)] is substantially greater than the eccentricity required to save the lengths of the retrograde arcs [Fig. 9(a)]. (Ptolemy says that the one is double the other, but this is only approximately so.)

Ptolemy's insight then consisted in realizing that he might preserve the correct spacing by leaving the center of uniform motion at the required distance from the Earth, and yet obtain correct regressions by placing the deferent's

center at half that distance.

Whether this was indeed the line of investigation that led Ptolemy to the equant cannot, of course, be proved. However, one additional piece of supporting evidence can be adduced. Immediately after completing the theory of longitudes for the five planets, Ptolemy shows (in *Almagest* XII, 1–6) how to use the new theory to calculate the lengths of the retrograde arcs for each planet. Ptolemy remarks that the lengths of the regressions should be shown to be in accord with those gotten from observations. Ptolemy calculates *the lengths of the retrograde arcs at apogee and at perigee for each planet*. Unfortunately, he does not report any observational values with which these theoretical results ought to be compared. Nevertheless, the contents of Book XII do demonstrate that matching the lengths of the retrograde arcs at apogee and perigee was an important consideration for Ptolemy. His language and his procedure imply that the results of these calculations were used as a test of his theory.

Finally, it will not be remiss to speculate on the planet that was occupying Ptolemy's attention when he came upon the equant. As we have seen, in his introduction to the theory of the superior planets, Ptolemy remarks that two independent methods of determining the eccentricity of the deferent in the intermediate model lead to two different results, the one being twice the other. According to one method, the eccentricity is chosen to save the spacing of the retrograde arcs about the zodiac. According to the other, the eccentricity is chosen to save the lengths of the retrograde arcs at apogee and at perigee. Now, in the case of Jupiter, the retrograde arcs at apogee and at perigee are almost identical, so that a calculation of the eccentricity from these data is not actually possible. Again, for Saturn, the case is hardly better. Only in the case of Mars is the difference between the lengths of the retrograde arcs so large that it would immediately suggest the use of these lengths in a derivation of the eccentricity. It seems most likely, then, that Ptolemy was led to the discovery of the equant while grappling with Mars, the planet that, fourteen centuries later, was to occupy the attention of Kepler.

¹R. Thiel, *And There Was Light*, translated by R. Winston and C. Winston (Mentor Books, New York, 1960), p. 60; published in Germany as *Und Es Ward Licht* (Rowohlt Verlag, Hamburg, 1956).

²The early history of the solar theory is not fully known. Although the eccentric orbit is usually associated with Hipparchos' name, the model may be somewhat older. Apollonios of Perge (c. 225 B.C.) is known to have experimented with eccentrics and epicycles and to have proved theorems concerning them. As the solar inequality was already known in Apollonios' time, it is quite likely that he attempted to explain this by means of the theoretical tools with which he is known to have been familiar. Thus the solar theory of Hipparchos may actually be due to Apollonios. Whether Apollonios went so far as to derive numerical parameters for the sun's orbit from observations is not known. It is possible that this step was first taken by Hipparchos. This seems to be another illustration of the gradual progression of Greek astronomical thought from qualitative physical explanation to quantitative mathematical theory.

³How can a planet, like Jupiter, which behaves according to the Ptolemaic model of Fig. 1, have retrograde arcs of uniform size? The following explanation may be of some help.

The retrograde arc tends to vary in size around the zodiac for two different reasons. First, the distance of the epicycle from the Earth varies from a maximum at apogee to a minimum at perigee. The epicycle will therefore look smaller at apogee and larger at perigee. Small epicycles

tend to produce small retrograde arcs. This distance effect therefore tends to make the retrograde arc smaller at apogee. But a second effect is operating simultaneously: the epicycle's center does not travel at a constant speed around the deferent. The motion is slower at apogee and faster at perigee. Now, retrograde motion of the planet occurs when the backwards motion of the planet on the epicycle overcomes the forward motion of the epicycle around the deferent. If the motion of the epicycle's center is slow, it will be more easily overcome, and a longer retrograde arc will result. Thus this speed effect tends to produce longer retrograde arcs at apogee.

To sum up: the distance effect acting alone would produce *smaller* retrograde arcs at apogee; but the speed effect acting alone would produce *larger* retrograde arcs at apogee. Which effect predominates depends on the parameters of the individual planet. In the case of Jupiter, the two effects cancel, so that the retrograde arc is of nearly constant size.

In the case of Saturn, the distance effect very slightly predominates, so that the retrograde arc at apogee is a tiny bit shorter than the retrograde arc at perigee. (Note that this is the reverse of the case with Mars: Mars has its longest retrograde arcs at apogee.)

⁴C. Ptolemy, *Syntaxis Mathematica* (or the *Almagest*). Greek text edited by J. L. Heiberg, *Claudii Ptolemaei Opera quae exstant omnia*, Vols. I and II (Teubner, Leipzig, 1898, 1903). English translations: R. Catesby Taliaferro, in *Great Books of the Western World* (Encyclopædia Britannica, Chicago, 1952), Vol. 16, pp. 1–478; and G. J. Toomer, *Ptolemy's Almagest* (Duckworth, London, 1984). Most of Ptolemy's scattered references to the history of the planetary theory are found in *Almagest* IX, 2 and XII, 1. Taliaferro's translation, while usable for some purposes, has de-

finitely been replaced by Toomer's. The senses of the two translations differ rather markedly in several brief passages in *Almagest* IX, 2.

⁵This point has been made, e.g., by J. L. E. Dreyer, *History of the Planetary Systems from Thales to Kepler* (Cambridge University, Cambridge, 1906); reprinted as *A History of Astronomy from Thales to Kepler* (Dover, New York, 1953), p. 166.

⁶Pliny, *Natural History*, II, 63–64. Latin text edited and English translation by H. Rackham *et al.* (Loeb Classical Library, 10 vols.) (Harvard University, Cambridge; William Heinemann, London; 1938–1962), Vol. I, pp. 210–213.

⁷There is a vestige of intermediate model B in Ptolemy's treatment of the superior planets. In the case of the final model (with a separate equant point) the calculation of the eccentricity from three oppositions is extremely difficult. Therefore Ptolemy begins by supposing that the equant point coincides with the center of the deferent. That is, he temporarily assumes intermediate model B. After obtaining an approximate value of the eccentricity from this calculation, he calculates the correction to the eccentricity produced by passing over to the final model in which the equant and the center of the deferent are separate (*Almagest* X, 7; XI, 1; XI, 5). In the first stage of the calculation we see, no doubt, the traditional method of determining the eccentricity in the intermediate model that Ptolemy had received from his predecessors.

⁸A brief discussion of several possible explanations, different from the one to be presented here, is given by O. Pedersen, *A Survey of the "Almagest"* (Odense University, Odense, 1974), pp. 277–279, 306–307.

⁹This is the opinion of O. Neugebauer, *A History of Ancient Mathematical Astronomy* (Springer, New York, 1975), Part I, p. 155.

¹⁰Pedersen, Ref. 8, p. 307.

Three perspectives on Schrödinger's cat

J. G. Loeser^{a)}

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

(Received 22 December 1983; accepted for publication 20 March 1984)

The resolutions to the paradox of Schrödinger's cat engendered by three major perspectives on quantum mechanics—quantum logic, the many worlds interpretation, and the Copenhagen interpretation—are discussed. The core of each perspective is revealed in the process. Each resolution to Schrödinger's cat has a counterpart which solves the logical paradox known as the unexpected hanging. The rather striking parallels serve to emphasize several points. First, the quantum paradox is an epistemological and not a physical one. Second, the three interpretations mentioned, unlike some others, are faithful to this understanding. And third, the pitfalls in these two paradoxes have a somewhat universal character.

I. PARADOXES IN PHYSICS

A paradox is an apparently acceptable argument leading to an absurd conclusion. In physics, one can distinguish two ways in which a paradox can arise. First, there is the possibility of an invalid line of reasoning. The paradox is then said to be due to a fallacy *in voce* (in discourse). Second, there is the possibility that some underlying assumption, like an accepted physical theory, is at fault. In this case the paradox is said to result from a fallacy *in re* (in the material). In each case, what renders the situation paradoxical is that it is not obvious just where the error lies, although it is usually clear to which category the paradox belongs.

An example of a paradox *in voce* is the pole and barn paradox.¹ This concerns a superman pole vaulter, his pole, and a barn slightly shorter than the pole. Since according to special relativity a quickly moving object appears contracted, if the vaulter runs with his pole at relativistic speed through the barn, they will disappear inside for an instant. From his point of view, however, this is absurd, since it is the barn that appears contracted. It is often posed as an instructive exercise to show how the contradiction is only apparent, and vanishes once it is recognized that whether one event (the disappearance of the back of the pole into the barn) comes before or after another (the emergence of the pole from the barn) can depend on the frame of reference.

A simple paradox *in re* is Olbers's starlight paradox.²